

Aspects of voice irregularity measurement in connected speech

Adrian Fourcin*

Emeritus Professor University College London

email

a.fourcin@ucl.ac.uk

based on:

Fourcin, A. Aspects of Voice Irregularity Measurement in Connected Speech
Folia Phoniatr Logop 2009;61:126–136 DOI: 10.1159/000219948

Voice irregularity measurement in connected speech

Summary

Applications of the use of connected speech material for the objective assessment of two primary physical aspects of voice quality are described and discussed. Simple auditory perceptual criteria are employed to guide the choice of analysis parameters for the physical correlate of pitch, and their utility is investigated by the measurement of the characteristics of particular examples of the normal speaking voice. This approach is extended to the measurement of vocal fold contact phase control in connected speech and both techniques are applied to pathological voice data.

Voice irregularity measurement in connected speech

Aspects of voice irregularity measurement in connected speech

Adrian Fourcin
Emeritus Professor University College London
a.fourcin@ucl.ac.uk

Introduction

Although we use voice as a dynamic time varying source of vocal tract excitation for all speech communication, connected speech is not currently the preferred signal for objective measurement in the voice clinic. In principle, however, it ought to provide crucially important data in the definition of what makes speech normal and serve as a central reference in the understanding and management of abnormal spoken communication. In particular, the rapid successions of different articulatory controls that are essential for the production of efficient speech sound sequences, subject laryngeal phonatory mechanisms to transient inputs that continually test their function to a depth that is quite unattainable by the sole use of sustained vowels.

Figure 1 Micro intonational voice regularity and irregularity

The present brief discussion is concerned with the quantitative clinical use of vocal fold frequency measurement in connected speech. The essential aim is to explain the basis of a simple approach that, in the first instance, makes use of auditory criteria in the definition of physical differences between normal regularity and unacceptable irregularity in voice pitch. Figure 1 shows aspects of the voiced component of a single word spoken with a falling intonation that, initially, has a smoothly changing vocal fold closure repetition frequency. F_x in the figure refers to the “instantaneous” frequency of vocal fold vibration (in contradistinction to F_0 that is ordinarily obtained from temporal windowing). The final part of the utterance, however, has obvious markedly irregular variations in cycle to cycle frequency (“pitch”), peak amplitudes and also contact phase durations. All of these factors, together with others that are not so readily visible, may contribute in their different ways and different degrees to the percept of roughness that ordinarily pervades our auditory appreciation of day to day speech. Only pitch variability, however, has begun to be adequately linked to psychophysical and psychophonetic measurements.

Voice irregularity measurement in connected speech

Table 1 Just detectable audible differences

Voice pitch

Table 1 gives a summary of the pitch perceptual sensitivities that are related to two broad classes of sounds — sustained and varying. The analysis of the physical attributes of steady vowels provides a mainstay for current laryngological voice assessment and it is important to note in the present context that the normal speaker's control of sustained voice pitch precisely matches the normal listener's auditory ability to discriminate voice pitch changes [3,4,5,6,7,8]. For sustained vowel assessment, this link between production and perception in respect of this dominant aspect of voice quality is an intrinsic supportive aspect of present clinical practice.

The values of DL (difference limen \equiv just noticeable difference) for the perception of voice pitch changes in connected speech are so markedly different from those found for sustained vowel perception, however, that a radically different auditory perceptual processing mechanism must be supposed to exist in our ordinary day to day use of spoken voice — both as listeners and also as speakers. The greater auditory integration times needed for DLs in singing, for example, would not enable the rapid processing required for normal speech comprehension. Although the DLs for sustained vowel pitch perception are not explicitly taken into account in voice analysis algorithms [9], they are attractive in providing a possible guide to the definition of protocols for the quantitative analysis of voice in connected speech. Spoken voice variability typically occurs often on a cycle to cycle basis within a speech sound and there are no extant perceptual measurements that provide a means of estimating their perceptual importance. The normally hearing speaker, however, continually monitors the nature of his/her speech and uses auditory feedback to signal the need for correction. It appears that it may be profitable to make use of an aspect of this process in physical measurement. In the limit, the auditory processing of the client and clinician is the final arbiter of unacceptable variability.

Before this can be done, however, there is a basic difficulty arising from the intrinsic nature of voice irregularity that must be clarified. The classic measurements made of perceptible pitch changes that are shown in Table 1 are all based on the perceptual discrimination of whole pitch-contours. The problem to be solved when dealing with the analysis of vocal frequency changes in ordinary speech arises from the need to provide an objective frame that enables changes from moment to moment in a given single utterance to be quantified.

Voice irregularity measurement in connected speech

Figure 2a A perceptual grid with superimposed notional production

Variability measurement

If the timed closure markers of Figure 1 are used to produce a cycle by cycle string of estimates of successive vocal fold vibrational frequencies (F_x) then these can be used to show and to measure cycle to cycle variability in a way that is relatable to pitch. In Figure 2 this has been done in a purely illustrative way using a grid of spaced lines that are separated by distances that are notionally related to DL pitch differences that apply to connected speech. This construction of fixed DL cells does not allow for the variation in DL values that may occur as a function of frequency and it is intended only to apply to the mode of voice frequency control that is associated with conversational speech. The singing voice, for example, would require more closely spaced DL bin lines in its analysis in order to cater for the greater precision of pitch control in production and in perceptual monitoring.

The variability that is potentially open to measurement is one that only extends over adjacent cycle pairs and, within each cycle of each pair, it is only necessary for the “instantaneous” frequency, F_x , of that cycle to fall within a bin for it to be included. Figure 2 has a notionally well controlled sample of voice pitch data superimposed on its grid lines — with variability in cycle to cycle variability that never exceeds 2 DLs. The central diagonal core of this crossplot (CF_x) contains only those cycle pairs whose members differ in F_x by less than one DL. The real test of this approach, in general and at this level of detail, however, can only properly come from its use with real speech data.

Figure 2b Actual production on a perceptual grid

Figure 2b has the same basic construction as that of Figure 2a but now the grid lines are spaced by half tone (6% F_x) intervals and the superimposed data has been derived from a recording of a read text by a speaker who is used to taking care of her voice. This association of an actual voice recording analysis with a perceptual grid gives a correspondance between production and perception that is not likely to be a matter of coincidence. It enables a simple metric of irregularity to be used that can be employed for the study of pitch control and vocal fold closure timing irregularity (see F_x in Figure 1) in the connected speech of both normal and pathological voices. When the separation between successive vocal fold instantaneous frequencies (F_x) exceeds two semitones then the corresponding pair of values falls into a cell that is outside the main central diagonal of this figure and the total of all of these occurrences can be used as a simple measure of irregularity, termed IF_x and expressed as a percentage of the total number of cycles in the whole sample. IF_x is 3.2% for this 2minute sample of a read text, so 96.8% of the

Voice irregularity measurement in connected speech

vocal fold cycles in the whole recording are associated with a cycle to cycle variability less than or equal to two semitones. Although the rationale basic to this measurement method has evolved interactively with Laryngologists, Speech and Language Therapists, Speech Scientists and Phoneticians over a number of decades and the implementation that has just been described has been in clinical use for more than two decades its rationale has not previously been spelled out. It has recently been the object of an excellent normative study by Pr Ptok of Hannover [11]. The approach is of use in the study both of aspects of normal speech production and also of the pathological voice.

Figure 3 Cfx examples for different normal speaker groups

Figure 3c shows the average Cfx crossplot for a particular group of fourteen English Speech and Language Therapy students. Their work has not required that they acquire exceptional speaking skills, and, paradoxically they are not given any special voice productive training. They are concerned with effective speech communication and are not called on to use expressive voice. Their mean vocal fold frequencies are similar and so also are their closure irregularities; the mean IFx of 6% is low and indicative of low closure timing irregularity.

Figure 3a is based on the analysis of the texts read by a quite different but similar age group of women Dramatic Arts students. A major objective of their training is concerned with voice production and the acquisition and control of expressive voice. It is immediately obvious from Figure 3a that this group has a greater voice frequency range than for 3c and it is also the case that their average irregularity score, IFX, of 11% is greater than that for the speech & language therapist group. One important aspect of voice control is related to voice quality, as well as voice frequency range, and an ingredient of contrastive phonation types is closely linked in English to the use of creaky voice. The ~ two minute read passage used by these students was designed so that there was no opportunity given for the use of "character" voices but, even so their training, after two years, has set them apart from the therapy students. The male dramatic art students have an IFx score of ~11% in figure 3b, and this also at first sight is surprising since it is higher than may be found for some other young men. These measurements are part of an ongoing study and although it is premature to draw firm conclusions, it is possible that the daily emphasis given to voice control has made these speakers, as for their women colleagues, use more creak than their contemporaries. It does not follow that either of these student groups has developed voice pathology as a result of their training [10] since the structure of their Cfx plots is not abnormal.

A rough rule of thumb has emerged from the clinical use of the Cfx plot

Voice irregularity measurement in connected speech

measurements that gives as a first guideline that an IFx of 10% is at an approximate borderline between good regular phonation (see also the quantitative normative work by Pr Ptok [11]). Voice pathologies are likely to be linked to much higher values of IFx and irregular CFx plots.

Figure 3d shows two data representations that are related to Figure 3b. The dotted, DFx1, outside plot makes use of all the Fx samples in the voice sample to show the range and way that their frequencies are distributed whilst the inner DFx2 plot only shows the distribution of successive Fx samples that occur in pairs falling within a single DL analysis bin. DFx2 corresponds to the diagonal bin of the CFx plot and the Coherence value is the ratio of these “regular” cycles to the size of the whole sample.

Figure 4 Pathological voice examples of CFx analysis

Figure 4 shows two sets of CFx crossplots for the comparison between the pre-operative and post-operative vocal fold closure timing regularity analyses for two phonosurgical patients. Figure 4a has four main lobes, two on the main diagonal at 170Hz and the main concentration at 300Hz. It is possible immediately to infer that, for a 35 year old woman, the voice will present as being unnaturally childish. The presence of the side lobes at [150,300Hz] and [300,150Hz] on either side of the main diagonal indicates a high probability of diplophonic octave saccades that will give a strong sensation of roughness. The second lobe in the main diagonal, at 170Hz, is within a more normal pitch range for the speaker but once more is linked to symmetrical side lobes at about [100,170Hz] and these also will give rise to perceived although lesser irregularity. Post-operatively, the vibrational CFx pattern is far less complex with very little irregular diplophonia at [205,120Hz] and [115,245Hz] and a main lobe now centred on 230Hz. This main lobe, however, is quite broad and extends for +\ - seven semitones and this indicates that, although there has been a considerable improvement with a change in IFx from 65% to 22% in closure timing regularity, the voice is still perceptually very rough. The data for the second set of Figures 4b and d relate to a quite different degree of dysphonia. There is a little change in the main modal frequency from 190Hz to 210Hz and initial IFx of 12% is already near to normal. The post-operative CFx plot shows a real improvement in regularity but in the auditory comparison of the samples there appears to be at least one other perceptual factor that should be taken into account and this is considered in the discussion of Figure 7.

Figure 5 Grade from GIRBAS evaluations and CFx measurements

The GRBAS scale was originally introduced for the perceptual grading of sustained vowel data [12] but it has been widely established for use in the

Voice irregularity measurement in connected speech

voice clinic with connected speech. A main objective of a study organised by Pr Dejonckere [13] was to calibrate an extended version of the scale, GIRBAS, and compare its findings with objective analysis by the cooperative work of pairs of clinicians in five institutes in three European countries. Two types of spoken material were used, for the objective analyses, the sustained vowel [ɑ:] and a short sentence (continually voiced and with neither fricatives nor voiceless plosives).

The subjective GIRBAS appraisal was based on a much longer sample of running speech for each patient and recorded from the consultation. In one centre, in Paris, Dr Fresnel simultaneously made CFx analyses during the recording of the consultation conversation. The results of the IFx values from these analyses are shown in Figures 5a and b, plotted together with her **G** ratings for each patient.

In Figure 5a the lowest value of **G (Grade)** (the most reliable of the GIBRAS grades) has been plotted against the associated value of IFx. This provides a way of selecting 17 patients for whom it is highly probable that Dr Fresnel made the grading on the basis of purely temporal vocal fold vibrational irregularity. There is, in this graph, a quite well defined trend to assign higher **G** values to recordings with greater IFx percentages.

Figure 5b has been plotted using the **G** and IFx% data to include the further 14 patients previously eliminated. To differing degrees these additional patients are not adequately represented, in the present objective terms, by the value of their associated IFx% measurements. In some cases Figure 5b shows that it is possible for a patient's voice to have a very low value of IFx but to be assigned high value **G**. Subsequent to this earlier work, voice clinicians (Mr McGlashan in Nottingham and Pr Ptok in Hannover) have observed that considerable voice quality improvement can result from treatment (phonosurgery and electrostimulation respectively) even though IFx% values pre and post-treatment were essentially the same. A major difference between the pre- and post-treatment objectively measured voice qualities for these patients appeared in the contact phase analyses of the Qx measurement shown in figure 1 [2] rather than in vocal fold vibratory closure epoch timing control. The tentative conclusions to be drawn in regard to the data presentations of Figure 5 is that a little under half of Dr Fresnel's group had voice quality characteristics that had both closure timing as well as other factors, particularly perhaps contact phase, contributing to their dysphonias. If contact phase is an important descriptor of voice quality then it must be a feature that is perceived by and productively controlled by normal speakers. In the context of the present work, this can be examined by making use of Qx crossplot analyses.

Voice irregularity measurement in connected speech

Figure 6 Contact phase calibration

Figures 6a and 6c show the contact phase cross plots, CQx, for the two groups of dramatic arts students for whom the CFx analyses were plotted in Figure 3. CQx is constructed in the same way as CFx with the important difference that the analysis bins are entirely heuristic and depend for their validity on the empirical evidence that normal speakers control their use of expressive vocal fold contact phase in ordinary speech in a consistent fashion. The two values of IQx shown in the figures are derived, in the same way as for IFx, with reference to a central core of bins, in CQx. The "DL" boundaries in CQx are also set by the definition of the bin size that is used. For these figures the bin size is 1% and the boundary beyond which irregularity is considered to exist is, as before, at +/- two bins or 2% on either side of the core diagonal. For the women students in Figure 6a the implication is that nearly 80% of their vocal fold vibrations are within an IQx% variability of ~20%; the Speech and Language Therapy students, however, have an average IQ of ~30%. For the men the corresponding coverage is about 65% for an IQx% of 35%. Figures 6b and 6d illustrate an important gender difference in the control of Qx between young men and young women. Men tend to have the same average contact phase control across their normal voice speaking frequency range. Many (although not all) young women tend to the production of smaller contact phase ratio in the upper part of their voice register [14]; this is shown here by the downward gradient of the women's average QxFx distribution. It is worth noting, however, that all these dramatic arts students control their overall use of voice timing in connected speech in the same way, so that voice takes up almost exactly 50% of the total speaking time. The other half of the total speaking time is taken up by voiceless fricatives and silences (an effect that has also been observed for other languages than English [2]).

Figure 7 Pathological voice examples of CQx analysis

Applications of this CQx analysis to the pathological voice data of Figure 4 are shown in Figure 7. A comparison of Figures 4a & c with Figures 7a & c shows that there is an obvious possibility that both closure timing and contact phase timing regularity changes have contributed to this patient's voice improvement. The phonosurgical intervention has reduced IQx from 87% to 53%, but still greater than the 20 to 30% for the women students. Figures 7b & d also show an improvement in Qx control following phonosurgery, with a final value of IQx% that is closer to what would normally be expected (and smaller than the IQx% of Figure 6c). Mr McGlashan has applied the Voice Handicap Index assessment technique to the assessment of patient voices and finds a striking correlation with changes in IQx and these brief particular examples also indicate that contact phase control is exerted by the speaker

Voice irregularity measurement in connected speech

and may, in consequence, be linked to voice quality improvement perceived by the listener.

Discussion

Connected speech not only has obvious relevance to voice use but is also capable of being used as the basis for objective voice quality analysis. At present EGG/Lx electrolaryngograph techniques provide an effective basis for signal processing but the ear's auditory analysis will eventually be sufficiently well modelled so that acoustic data alone may be sufficient to enable the physical application of perceptual processing.

A special emphasis has been given here to the use of auditory psychophonic measurements in defining analysis criteria and this is likely to be merely the beginning of further developments that will make more use of accumulating neurophysiological observations that relate to the mechanisms of temporal pitch processing [15]. The remarkable human auditory ability to adjust its effective integration time to suit the analysis of widely different acoustic inputs, so that sung voice pitch may be heard with tenfold greater precision than that for connected speech (Table 1), will have to be incorporated explicitly in the voice analysis algorithms of the future.

It was not evident that contact phase variation might be such an important component of connected speech voice quality until clinical observation resulted in data that could only be given this explanation, and the fact that speech is controlled in this regard implies also that interactive voice therapy procedures may benefit from appropriate related biofeedback technology that enables training to be monitored by analysis of this voice dimension.

Finally, this general progress can only be achieved as the result of the type of concerted activity that was mentioned earlier in connection with the assessment of the GIRBAS scale. Multi-centre, poly language enabling research can put the future efficient use of connected speech into the category of a reliable and simple calibrated every day tool.

Acknowledgements

This work would not have been feasible without the critical input and the help of Evelyn Abberton and data and discussion, from and with: Elisabeth Fresnel, Julian McGlashan, Martin Ptok, Romain Pérouse, Bruno Coulombeau, the help of staff and students at RADA, LAMDA, University College London; and both hardware and software from Xinghui Hu (Laryngograph.com) and it would not have appeared without the incentive and forum for discussion provided by Philippe Dejonckere and other COST 2103 colleagues.

Voice irregularity measurement in connected speech

References

- 1 Fabre,P.Un procédé électrique percutané d'inscription de l'accolement glottique au cours de la phonation. Bull. Nat. Méd 1957, 141, 66-99
- 2 Fourcin A, Abberton E, Hearing and phonetic criteria in voice measurement: Clinical applications; Logopedics Phoniatrics Vocology. 2008; 33: 35_48
- 3 Moore BCJ, Glasberg BR, Peters RW. Relative dominance of individual partials in determining the pitch of complex tones. J Acoust Soc Am. 1985; 75: 550_61.
- 4 Nelson, D.A., Stanton, M. E., and Freyman R. L., 1983, "A general equation describing frequency discrimination as function of frequency and sensation level" J.A.S.A. 73, 2117-2123
- 5 Vance, T.F.,1914, see page 340 in E.G. Boring, "Sensation and Perception in the History of Experimental Psychology", Appleton-Century-Crofts Inc. New York 1942
- 6 Wier, C.C., Jesteadt, W., Green, D.M., 1977, "Frequency discrimination as a function of frequency and sensation level", J.A.S.A. 61, 178-184
- 7 Klatt, D. H. 1973, "Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception", J.A.S.A. 53, 8-16
- 8 't Hart, J.,1981,"Differential sensitivity to pitch distance, particularly in speech" J.A.S.A. 65, 811-821
- 9 Buder EJ. Acoustic analysis of voice quality. In: Kent RJ, Ball MJ, editors. Voice quality measurement. San Diego: Singular (Thompson Learning); 2000.
- 10 Walzak P, McCabe P, Madill C, Sheard C., Acoustic Changes in Student Actors' Voices After 12 Months of Training. J. Voice 2008, 22(3), 257-384
- 11 Ptok M, Iven C, Jessen M, Schemmle C. Objektiv gemessene Stimmlippenschwingungsirregularität vs. subjektiver Eindruck der Rauigkeit. Zeitschrift HNO. 2006; 54: 132_6.
- 12 Hirano M. Clinical examination of voice. New York: Springer-Verlag; 1981.
- 13 Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier-Buchman L, Millet B. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Revue de Laryngologie Otologie Rhinologie 1996; 117(3):219-224
- 14 Abberton ER. Aspects of Voice Quality in women 2002. Measurement of Speech Sound Data and its practical application 128-131 Proceedings of the First Conference on Speech Science, Seoul Korea
- 15 Sayles, M., and Winter, I.: Ambiguous Pitch and the Temporal Representation of Inharmonic Iterated Rippled Noise in the Ventral Cochlear Nucleus. The Journal of Neuroscience, 2008, 28(46):11925-11938;

Voice irregularity measurement in connected speech

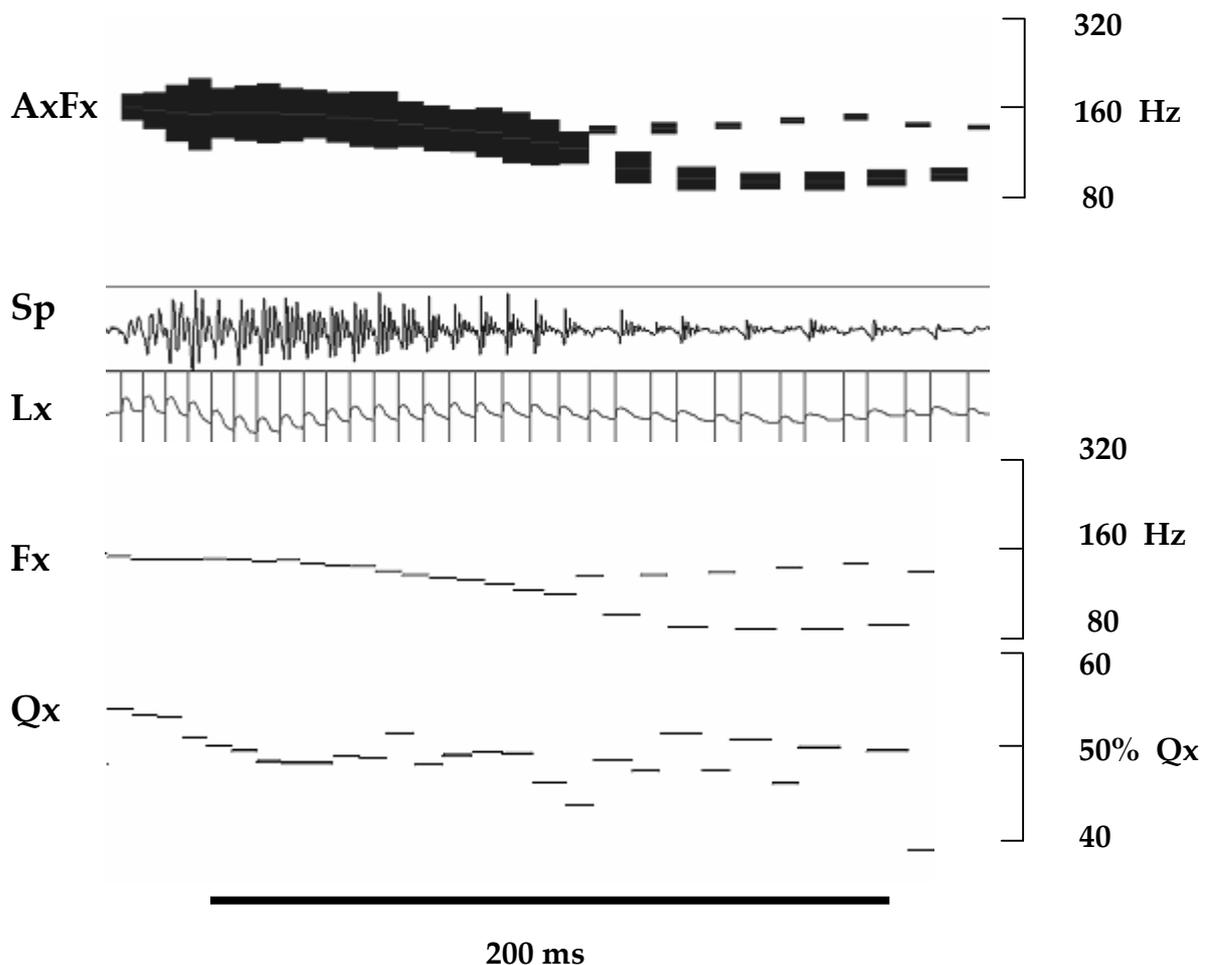


Figure 1

Micro-intonational voice regularity and irregularity

- Sp** Speech pressure waveform – compression positive
- Lx** EGG waveform with ELG polarity – contact positive [1, 2]
- Tx** Vocal fold cycle – duration between successive Lx markers
- Ax** Sp positive pressure peak after Lx contact marker
- Fx** “Instantaneous Frequency” – derived from Tx
- AxFx** composite “intonation” trace made by width modulating the cycle by cycle vocal fold instantaneous frequency trace with the corresponding Sp peak values
- Qx** % ratio to Tx of Lx contact duration – 70% below waveform peak

The sampling rate used to determine Tx and Fx is 1MHz

Voice irregularity measurement in connected speech

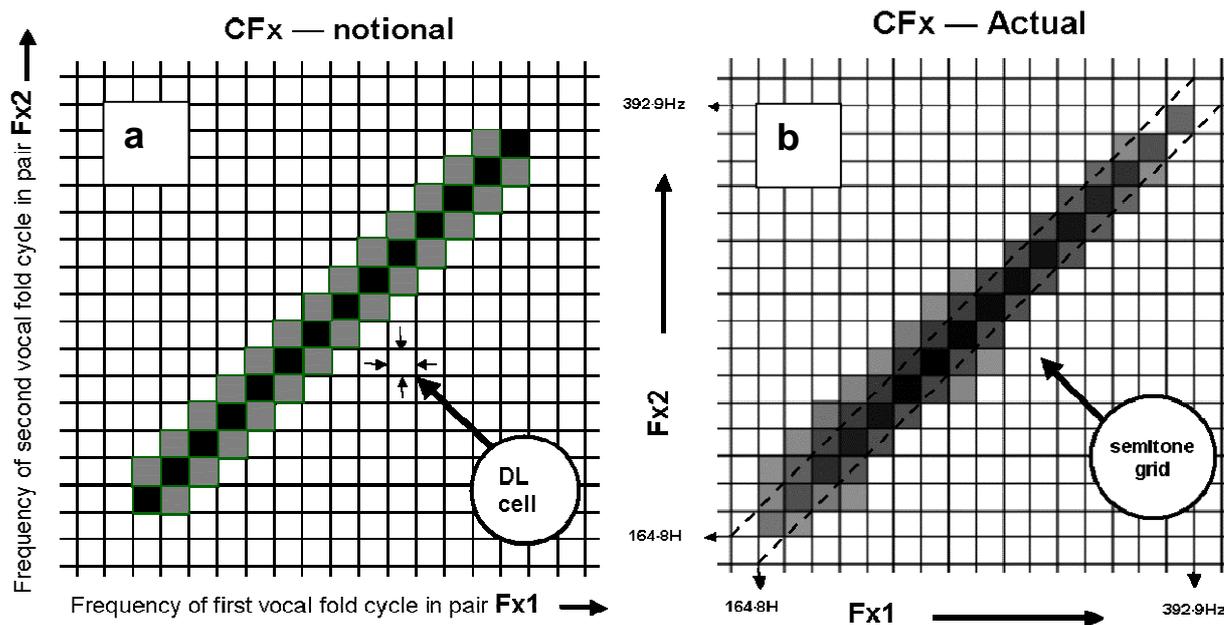


Figure 2

a A perceptual grid with superimposed notional production

The spacings between adjacent vertical lines and those between adjacent horizontal lines are intended to correspond to a just auditorily distinguishable difference in voice frequency (DL). The black filled bins show the core result if a voice sample with no discernable cycle to cycle variability were to be analysed in these terms on a cycle pair basis. The light gray bins relate to the occasions in a smooth Fx contour for which there are two DL changes in pitch within a successive pair of cycles.

b An actual production on a perceptual grid

Expanded CFx crossplot analysis of a two minute sample of normal speech, produced by a occupational voice user, with superimposed perceptual semitone grid and dashed reference double semitone DL lines.

15,250 vocal fold cycles have been analysed and 488 have fallen outside the cells marked by the dashed upper and lower DL lines giving an index of variability

Ix of 3.2% = 488/15,250

Voice irregularity measurement in connected speech

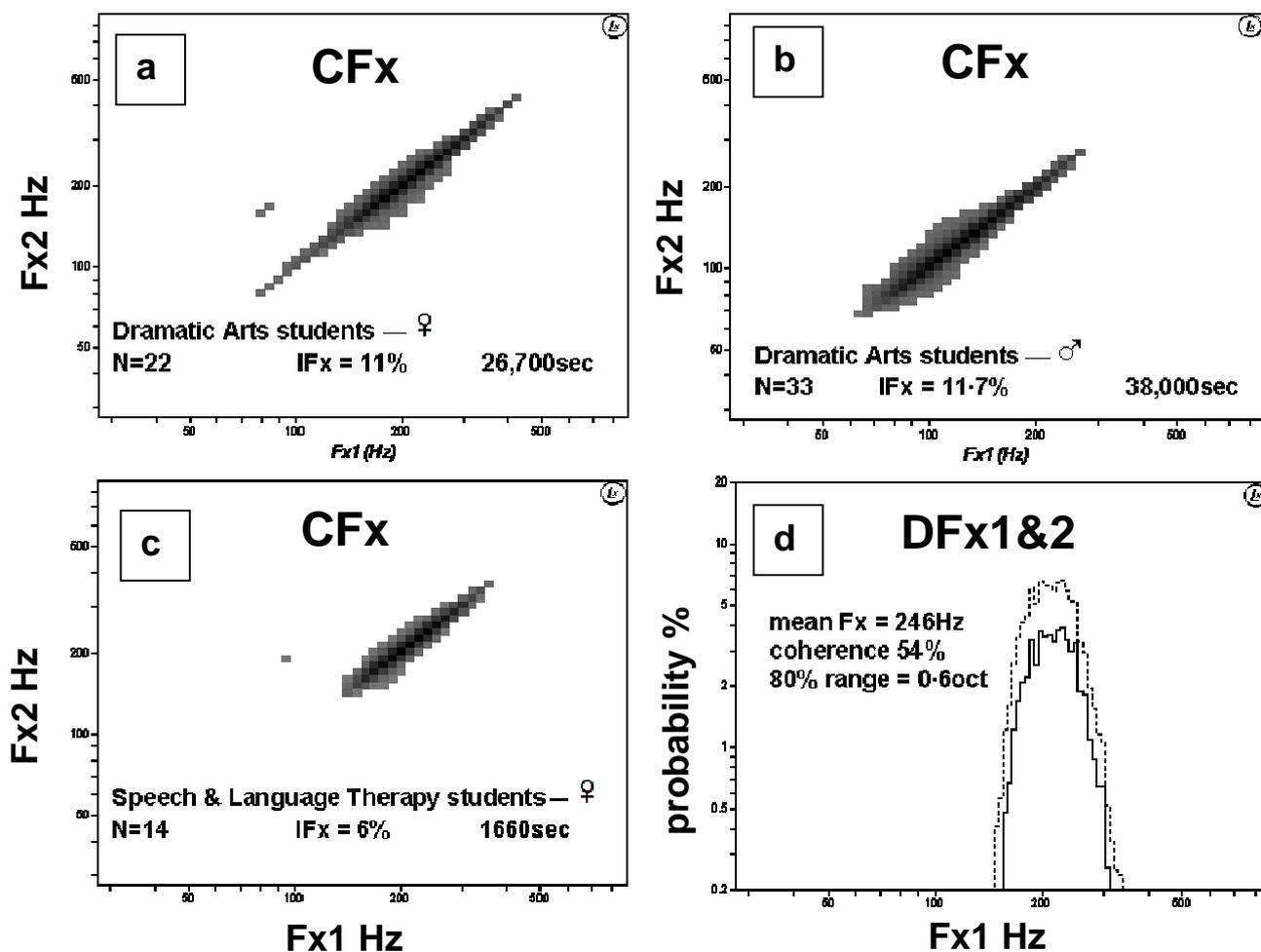


Figure 3

CFx examples for different normal speaker groups

For Figures a, b and c the speech and laryngograph data samples for each member of the three separate groups of speakers have been grouped together and analysed as single files. This technique gives an overview of the average irregularity that is characteristic of the speaker class.

Figure 3 d is based on the same data as that for Figure 3c and shows the probability distribution of Fx occurrences using 3% quarter tone wide bins. DFX1 gives only the probabilities of occurrence for every vocal fold cycle in the whole sample whilst DFX2 gives the probabilities for pairs of successive cycles falling within the same 3% Fx bin. DFX2 is linked to the centre diagonal of the corresponding CFX analysis of the data — but like DFX1 makes use of a smaller width.

[Both CFX and the DFX plots start at 27.5Hz (A0) and end at 880Hz (A5); #sec refers to the duration in seconds of the whole amalgamated sample].

Voice irregularity measurement in connected speech

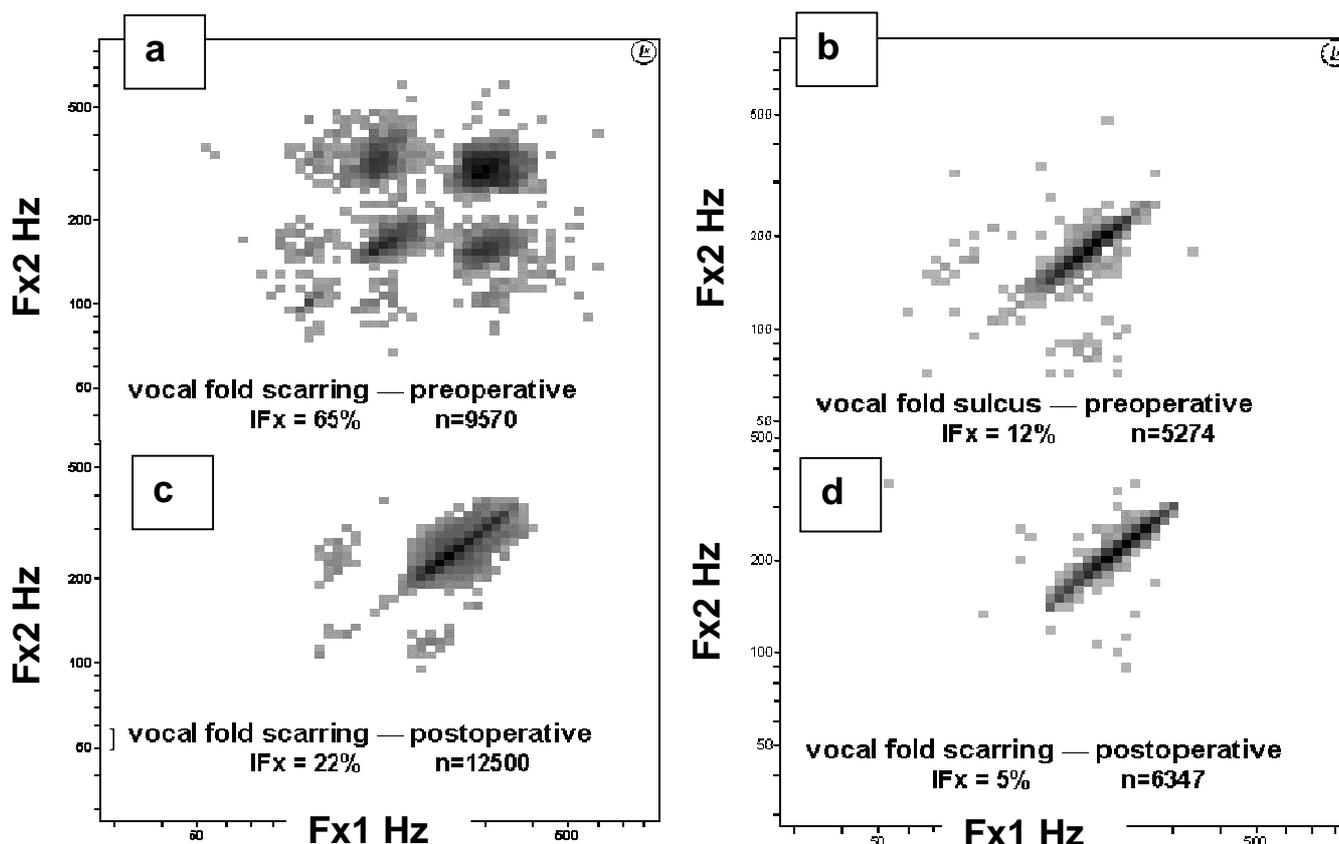


Figure 4

Pathological voice examples of CFx analysis

The first pair of examples on the left in Figures 4a and 4c concern a single patient whose primary voice quality disability arises from vocal fold closure timing, Fx, irregularity. In these cases the use of the I_{fx} index is of direct management utility and relates to the auditory impression of the associated voice quality.

There may, however, be additional factors that influence voice quality. Figure 5b is linked to a poorer voice quality than might be supposed from an I_{fx} of 12% and Figure 5d is associated with a change to almost normal voice quality. See Figure 8 and the associated text for further discussion.

(Data for Figures 5a & c by courtesy of Mr McGlashanFRCS, Nottingham and for Figures 5b & d from Dr Coulombeau and Dr Pérouse, Lyons)

Voice irregularity measurement in connected speech

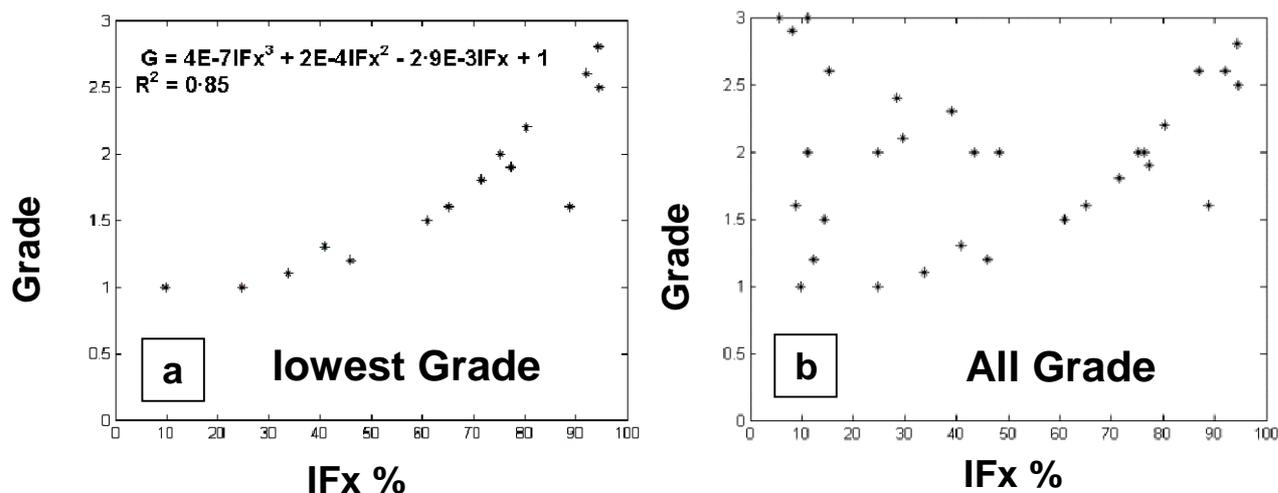


Figure 5

Grade from GIRBAS evaluations and CFx measurements

This comparison comes from Grade evaluations made by Dr Fresnel — as part of a study organised by Pr Dejonckere [10].

Assessments for 32 patients are shown. One half of the IFx values from CFx analyses lie on the lower border of the set of Grade values — Figure 5a. For these values of Grade the dysfunction seems likely to be purely related to closure irregularity. The inference from these results is that for the other half of the voice recordings included in Figure 5b, there are other sources of dysfunction that contribute to the Grade assessment.

([10] is the source for the GIRBAS study; [11] discusses other sources of irregularity and their measurement)

Voice irregularity measurement in connected speech

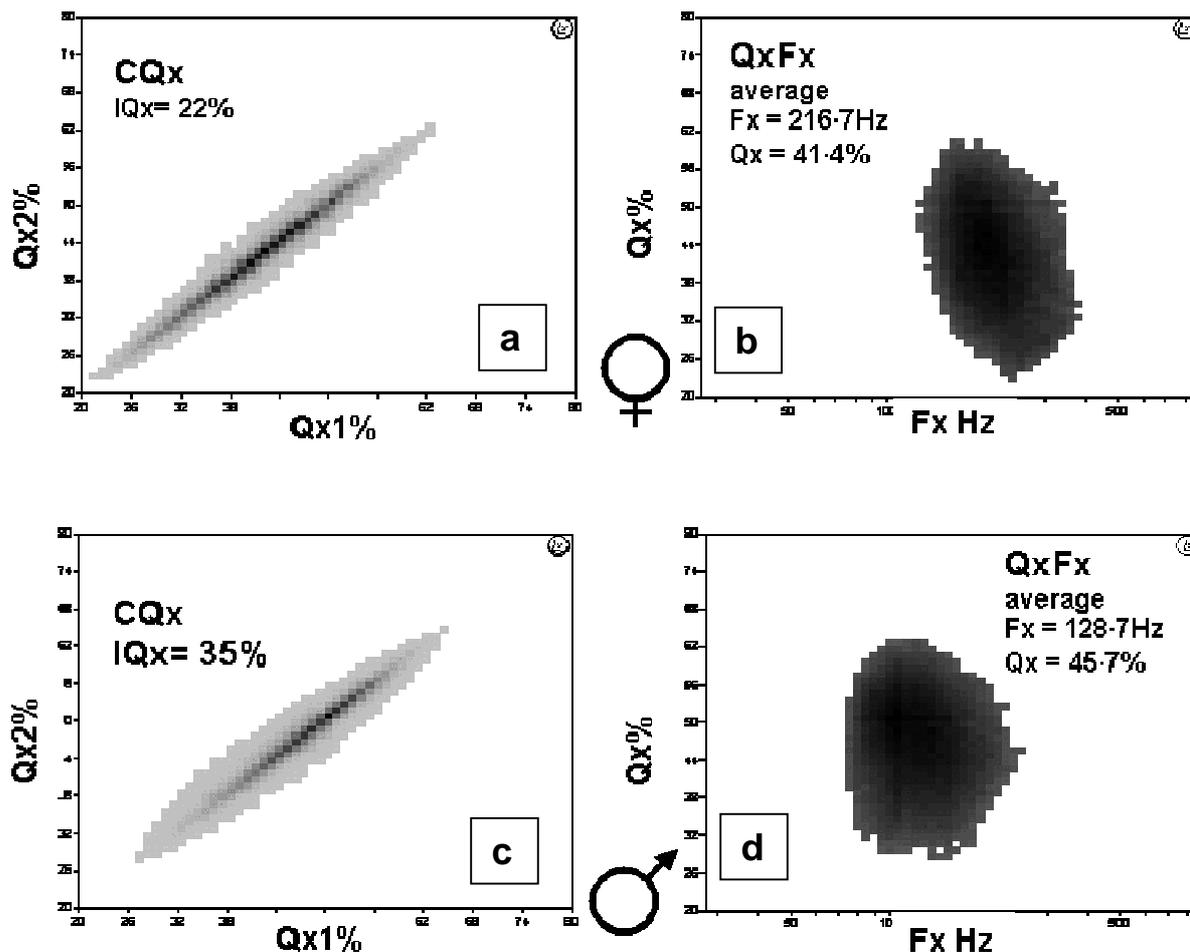


Figure 6

CQx and FxQx examples for two normal speaker groups

In figure 7 the data samples for each member of the separate groups of speakers have been grouped together and the resulting amalgamation has been analysed as a single file.

Figures a and c, show the paired cycle to cycle variabilities arising from changes in vocal fold contact phase duration. This aspect of voice production has proved to be of clinical importance and these samples give an indication of variation that might be found with normal larynges.

Figures b and d show how these men and women speakers, on average, control their contact phase, Q_x , as a function of vocal fold frequency, F_x . There is a distinct gender contrast. This is likely to prove to be language independent since it arises primarily from anatomical differences.

(The CQx and Qx plots start at $Q_x=20\%$ and end at $Q_x=80\%$ with 60 equally spaced 1% bins, whilst F_x , as before is in the range A0 to A5)

Voice irregularity measurement in connected speech

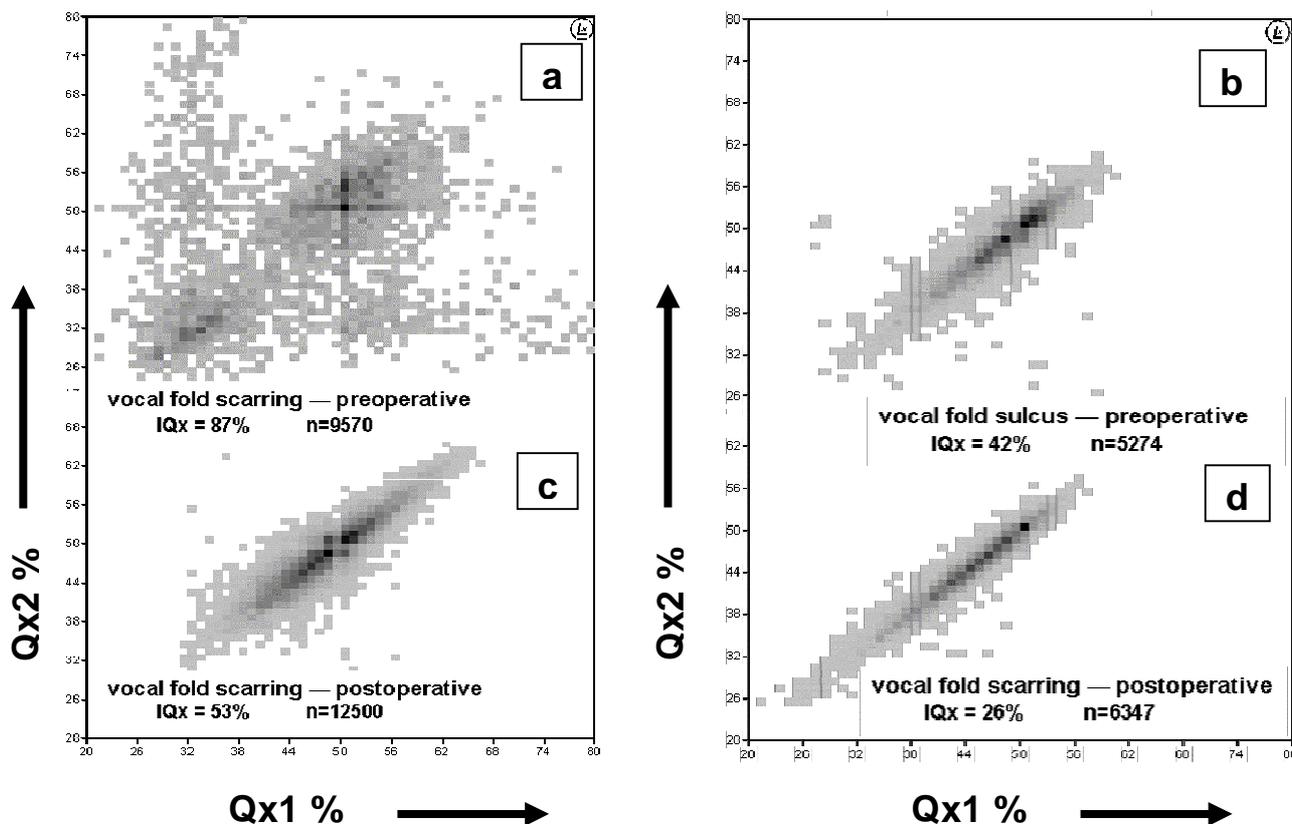


Figure 7

Pathological voice examples of CQx analysis

Figure 5a indicates that, for that particular data set of 32 patients, it was likely that half of the cases had purely temporal, Fx, vocal fold closing epoch variability. For the other cases included in Figure 5b, additional factors, including contact phase duration, were possibly involved. The crossplot analyses of Qx (CQx) that are shown here, give examples of the improvements in contact phase, Qx, regularity that may result from phonosurgical intervention; and that occur in addition to the improvements that are linked to closure epoch timing shown in the CFx crossplots of Figure 4 for the same voice patients.

Voice irregularity measurement in connected speech

Sound Type	pure tone	vowel		connected speech
		sustained	fall or rise	
frequency range	200 Hz to 1kHz	200 Hz to 600 Hz	100 Hz	60 Hz to 180 Hz
DL semitones	0.07 to 0.12	0.02 to 0.05	1.3	1 to 3.3
source	Nelson, et al, 1983 Vance,,1914,	Wier, et al, 1977 Moore, et al 1985,	Klatt, 1973	'tHaart, 1981

Table 1

Psychophysical measurements of just noticeable pitch differences (pitch difference limens) for sustained sounds and connected (synthetic) speech.